

To appear in The Astrophysical Journal

MASS SEGREGATION IN STAR CLUSTERS: ANALYTIC ESTIMATION OF THE TIMESCALE

HIDEAKI MOURI

*Meteorological Research Institute, Nagamine 1-1, Tsukuba 305-0052, Japan;
hmouri@mri-jma.go.jp*

and

YOSHIAKI TANIGUCHI

*Astronomical Institute, Graduate School of Science, Tohoku University, Aoba, Sendai 980-8578,
Japan; tani@astr.tohoku.ac.jp*

ABSTRACT

Mass segregation in a star cluster is studied in an analytical manner. We consider a two-component cluster, which consists of two types of stars with different masses. Plummer’s model is used for the initial condition. We trace the overall behaviors of the probability distribution functions of the two components and obtain the timescale of mass segregation as a simple function of the cluster parameters. The result is used to discuss the origin of a black hole with mass of $\gtrsim 10^3 M_\odot$ found in the starburst galaxy M82.

Subject headings: stellar dynamics — galaxies: star clusters — galaxies: starburst

1. INTRODUCTION

Within a star cluster, massive objects segregate into the cluster core as they lose kinetic energies to the less massive objects during approach toward energy equipartition. The result is an increase of the number density of the massive objects in the core. This phenomenon of mass segregation has long been known (Spitzer 1969; Binney & Tremaine 1987, p. 531) but has attracted new interests in recent years. The large number density induces successive mergers of those massive objects, which could lead to the formation of an exotic object such as a very massive black hole.

The important parameter of mass segregation is its timescale. In order to explain an observation in terms of mass segregation, the timescale has to be less than the age of the star cluster.

However, mass segregation has been studied only for specific cases using numerical methods, i.e., Fokker-Planck, Monte Carlo, and N -body simulations. The exceptional analytical works that

are available are those of Spitzer (1969) and Tremaine, Ostriker, & Spitzer (1975). Spitzer (1969) considered an isothermal cluster of uniform density and estimated the timescale of mass segregation as the timescale of energy equipartition between stars of different masses. This model is too idealized for a star cluster, which is not isothermal or of uniform density. Tremaine et al. (1975) studied a decaying circular motion of an object in the singular isothermal sphere. This is originally a model for a motion of a globular cluster around the center of a galaxy, and it is not a model for random motions of stars in a cluster. Therefore, an analytical study of a more appropriate situation is desirable.

We analytically study a two-component cluster, which consists of two masses m_0 and m ($m_0 < m$). The major component is the less massive stars. Their total mass M_0 dominates over the total mass M of the more massive stars, $M_0 \gg M$. Here the subscript 0 is used to indicate that the quantity is relevant to the major component. The cluster is spherically symmetric in all its properties and completely isotropic in the velocity space.

Since the timescale of mass segregation is expected to be less than the relaxation timescale (Spitzer 1969), the star cluster is regarded as almost collisionless. The individual stars move almost freely in the mean potential. The gravitational encounters occur only in a stochastic manner. Our study is accordingly based on the overall behaviors of the probability distribution functions (PDFs) of the major and minor components. They are approximated by PDFs of completely collisionless systems. The minor-component PDF is from a parameterized family of steady-state PDFs for which the degree of mass segregation is allowed to change. The major-component PDF is kept the same (§2). These PDFs are used to obtain the energy loss per unit time from the minor component due to encounters with the major component (§3). The energy loss rate yields the rate of change of the minor-component PDF, which in turn yields the timescale of mass segregation (§4, eq. [26]).

We compare our result with those for relevant timescales, including the pioneering work of Spitzer (1969). We also use our result to discuss the origin of a black hole with mass of $\gtrsim 10^3 M_\odot$ found in a young star cluster of the starburst galaxy M82 (§5).

2. PROBABILITY DISTRIBUTION FUNCTIONS

The major component, i.e., the less massive stars, is assumed to follow Plummer’s model (polytrope with $n = 5$; Binney & Tremaine 1987, p. 223). This steady-state solution for a collisionless spherical system has a simple analytic form but well reproduces observations of some globular clusters. The PDF in the phase space is determined by two parameters, e.g., the total mass M_0 and the core radius r_c :

$$f_0 = \frac{32\sqrt{2}}{7\pi^2} \frac{3M_0}{4\pi r_c^3} \left(\frac{GM_0}{r_c} \right)^{-5} \left(\Psi_0 - \frac{v^2}{2} \right)^{7/2}, \quad (1)$$

with $f_0 = 0$ at $v \geq (2\Psi_0)^{1/2}$. Here G is the gravitational constant and v is the velocity. The relative potential Ψ_0 is a function of the radius r and defined as the negative of the gravitational potential

$\Phi_0 \leq 0$:

$$\Psi_0(r) = -\Phi_0(r) = \frac{GM_0}{r_c} \frac{1}{(1 + r^2/r_c^2)^{1/2}}. \quad (2)$$

The corresponding mass density ρ_0 is

$$\rho_0(r) = 4\pi \int_0^\infty v^2 f_0 dv = \frac{3M_0}{4\pi r_c^3} \frac{1}{(1 + r^2/r_c^2)^{5/2}}. \quad (3)$$

Here the integral $4\pi \int_0^\infty r^2 \rho_0 dr$ gives the total mass M_0 . The squared velocity dispersion σ_0^2 is

$$\sigma_0^2(r) = \frac{4\pi \int_0^\infty v^4 f_0 dv}{4\pi \int_0^\infty v^2 f_0 dv} = \frac{GM_0}{2r_c} \frac{1}{(1 + r^2/r_c^2)^{1/2}}. \quad (4)$$

The minor component, i.e., the more massive stars, is assumed to have no contribution to the gravitational field of the star cluster. Then the gravitational field is determined solely by the relative potential Ψ_0 of the major component. According to the Jeans theorem for a steady-state collisionless spherical system (Binney & Tremaine 1987, p. 221), any function of $\Psi_0 - v^2/2$ serves as the minor-component PDF. The PDF assumed in our study is from a function family that has a parameter $\beta \geq 1$:

$$f_\beta = \frac{32\sqrt{2}}{7\pi^2} F_1 F_2 \frac{3M}{4\pi r_c^3} \left(\frac{GM_0}{r_c} \right)^{-(7\beta+3)/2} \left(\Psi_0 - \frac{v^2}{2} \right)^{7\beta/2}, \quad (5)$$

with $f_\beta = 0$ at $v \geq (2\Psi_0)^{1/2}$. The factors F_1 and F_2 are functions of β defined as

$$\int_0^\infty \frac{x^2 dx}{(1 + x^2)^{(7\beta+3)/4}} = \frac{1}{3} \frac{1}{F_1(\beta)}, \quad (6)$$

and

$$\int_0^1 x^2 (1 - x^2)^{7\beta/2} dx = \frac{7\pi}{512} \frac{1}{F_2(\beta)}, \quad (7)$$

where $F_1 = F_2 = 1$ at $\beta = 1$. Their numerical values are shown in Table 1 and Figure 1. The corresponding mass density ρ_β is

$$\rho_\beta(r) = F_1 \frac{3M}{4\pi r_c^3} \frac{1}{(1 + r^2/r_c^2)^{(7\beta+3)/4}}, \quad (8)$$

with the total mass M . The squared velocity dispersion σ_β^2 is

$$\sigma_\beta^2(r) = \frac{12}{7\beta + 5} \frac{GM_0}{2r_c} \frac{1}{(1 + r^2/r_c^2)^{1/2}} \propto \sigma_0^2(r). \quad (9)$$

Figure 2 shows the radial profiles of the mass density and squared velocity dispersion for $\beta = 1, 10$, and 100. The total energy E_β is obtained by integrating $\rho_\beta(\sigma_\beta^2 + \Phi_0) = \rho_\beta(\sigma_\beta^2 - \Psi_0)$ over the entire volume:

$$E_\beta = -\frac{3\pi}{32} \frac{2(7\beta - 1)}{7\beta + 5} F_1 F_3 \frac{GM_0 M}{r_c}. \quad (10)$$

The factor F_3 is a function of β defined as

$$\int_0^\infty \frac{x^2 dx}{(1+x^2)^{(7\beta+5)/4}} = \frac{\pi}{16} F_3(\beta), \quad (11)$$

where $F_3 = 1$ at $\beta = 1$. We define the initial state as $f_{\beta=1}$ so that the radial profile of mass density and the velocity dispersion are the same as those of the major component. With an increase of the β value, the mass density of the minor component becomes more centrally concentrated than that of the major component. Simultaneously, the velocity dispersion of the minor component becomes smaller (Fig. 2). These behaviors represent mass segregation. The radial profile of velocity dispersion of the minor component remains the same as that of the major component. This property is essential to establishing the final state of mass segregation where no net energy exchange exists between the two components (§3). Thus, with a single parameter $\beta \geq 1$, the function family f_β provides the simplest model to trace the evolution of the minor component from the initial to final states of mass segregation. At an intermediate state, the minor-component PDF is set to be just one of the functions, $f = f_\beta$, or a linear combination of them, $f = \sum (M_\beta/M) f_\beta$ with $M = \sum M_\beta$ (§4).

3. ENERGY CHANGE RATE

Consider a single object in the minor component, which loses kinetic energy due to gravitational encounters with objects in the major component. If the velocity of the single object is v , the rate of change of kinetic energy per unit mass is obtained as

$$\left\langle \frac{\partial}{\partial t} \left(\frac{v^2}{2} \right) \right\rangle = 16\pi^2 G^2 \ln \Lambda \left(m_0 \int_v^\infty v_0 f_0 dv_0 - m \int_0^v \frac{v_0^2}{v} f_0 dv_0 \right), \quad (12)$$

with $f_0 = f_0(\Psi_0 - v_0^2/2)$ (Binney & Tremaine 1987, eq. [8-66]). Here $\langle \cdot \rangle$ denotes an ensemble average. The cutoff factor Λ is the ratio of the maximum and minimum impact parameters, b_{\max} and b_{\min} :

$$\Lambda = \frac{b_{\max}}{b_{\min}} = \frac{b_{\max} v_{\text{rel}}^2}{G(m+m_0)} \simeq \frac{m_0}{m+m_0} N_0. \quad (13)$$

Here v_{rel} is a typical relative velocity between objects in the cluster (Binney & Tremaine 1987, p. 511), and $N_0 = M_0/m_0$ is the total number of objects in the major component. We have assumed $v_{\text{rel}}^2 \simeq GM_0/b_{\max} = Gm_0 N_0/b_{\max}$.

By multiplying equation (12) by the minor-component PDF f and integrating over the v space, we obtain the rate of change of kinetic energy per unit volume of the minor component. For the PDF $f = f_\beta(\Psi_0 - v^2/2)$, the result is

$$\begin{aligned} \left\langle \frac{\partial}{\partial t} \left(\frac{\rho \sigma^2}{2} \right) \right\rangle &= -\frac{2^{19} \sqrt{2} 16!!}{49\pi 19!!} F_1 F_2 F_4 \frac{1}{(1+r^2/r_c^2)^{(7\beta+12)/4}} \\ &\quad \times G^2 \ln \Lambda \left(\frac{m}{7\beta+2} - \frac{m_0}{9} \right) \frac{3M_0}{4\pi r_c^3} \frac{3M}{4\pi r_c^3} \left(\frac{GM_0}{r_c} \right)^{-1/2}. \end{aligned} \quad (14)$$

The factor F_4 is a function of β defined as

$$\int_0^1 x^2 (1 - x^2)^{(7\beta+9)/2} dx = \frac{16!!}{19!!} F_4(\beta), \quad (15)$$

where $F_4 = 1$ at $\beta = 1$. By integrating equation (14) over the entire volume, we obtain the rate of change of total energy E of the minor component:

$$\begin{aligned} \frac{dE}{dt} = & -\frac{2^{21}\sqrt{2\pi}}{3 \times 5 \times 7^2} \frac{16!!}{19!!} \frac{\Gamma(13/4)}{\Gamma(15/4)} F_1 F_2 F_4 F_5 r_c^3 \\ & \times G^2 \ln \Lambda \left(\frac{m}{7\beta + 2} - \frac{m_0}{9} \right) \frac{3M_0}{4\pi r_c^3} \frac{3M}{4\pi r_c^3} \left(\frac{GM_0}{r_c} \right)^{-1/2}. \end{aligned} \quad (16)$$

The factor F_5 is a function of β defined as

$$\int_0^\infty \frac{x^2 dx}{(1 + x^2)^{(7\beta+12)/4}} = \frac{\sqrt{\pi}}{15} \frac{\Gamma(13/4)}{\Gamma(15/4)} F_5(\beta), \quad (17)$$

where $F_5 = 1$ at $\beta = 1$.

Recall that gravitational encounters in a star cluster are stochastic. The energy change rates (12) and (14) have been obtained as ensemble averages. They are not locally defined quantities. The energy change rate for the entire minor component (16) has been obtained by assuming $\langle dE/dt \rangle = dE/dt$ because the cluster contains many stars.

The first and second integrals in equation (12) represent heating and cooling by the major component, respectively. They yield the terms $m_0/9$ and $m/(7\beta + 2)$, respectively, in equations (14) and (16). When the two terms balance, the total energy of the minor component settles to an equilibrium, $dE/dt = 0$. Hence the β value corresponding to the final state of mass segregation is

$$\beta_{\text{fin}} = \frac{9m - 2m_0}{7m_0}. \quad (18)$$

This relation does not imply equipartition of kinetic energies, $m\sigma^2 = m_0\sigma_0^2$ (see eqs. [4] and [9]), because the major component is kept to follow Plummer’s nonthermal model. The condition for energy equipartition in an equilibrium state is that each of the two components has a thermal velocity distribution (Chapman & Cowling 1970, p. 80). A qualitatively consistent result for very large values of m/m_0 was obtained by Chatterjee, Hernquist, & Loeb (2002).¹

¹Chatterjee et al. obtained the squared velocity dispersion σ^2 at $r \rightarrow 0$ for a Brownian motion of a single massive object that is in equilibrium with the background stars. Their study is mathematically equivalent to ours because we have ignored encounters among objects in the minor component. The background stars were assumed to follow Plummer’s model as in our present study.

4. TIMESCALE OF MASS SEGREGATION

By using the energy change rate dE/dt in equation (16), we estimate the timescale of mass segregation τ_{ms} . This timescale is defined in terms of the secular increase of the mass of the minor component, i.e., the more massive stars, in the cluster core M_c :

$$\tau_{\text{ms}} = \left. \frac{M_c}{dM_c/dt} \right|_{t=0}, \quad (19)$$

with

$$M_c = 4\pi \int_0^{r_c} r^2 \rho dr = \frac{F_1 F_6}{2\sqrt{2}} M, \quad (20)$$

for $f = f_\beta$. The factor F_6 is a function of β defined as

$$\int_0^1 \frac{x^2 dx}{(1+x^2)^{(7\beta+3)/4}} = \frac{1}{6\sqrt{2}} F_6(\beta), \quad (21)$$

where $F_6 = 1$ at $\beta = 1$.

Since it is not analytically determined how the minor-component PDF changes in the course of mass segregation, we study two extreme cases. The one is that the minor-component PDF is characterized by a single β value, which increases from the initial value of 1 toward the final value β_{fin} , i.e., $f = f_\beta$ with $\beta = 1$ at $t = 0$ ($1 \leq \beta \leq \beta_{\text{fin}}$). Equations (6), (20), and (21) yield $dM_c/d\beta$, while equations (6), (10), and (11) yield $dE/d\beta$. Then dM_c/dt is obtained as $(dM_c/d\beta)(d\beta/dt) = (dM_c/d\beta)(dE/dt)/(dE/d\beta)$. Using their initial values, we obtain

$$\begin{aligned} \tau_{\text{ms}} = & \frac{3^3 \times 5 \times 7^2 \pi \sqrt{\pi} 19!! \Gamma(15/4)}{2^{23} \sqrt{2} 16!! \Gamma(13/4)} \frac{7 - 8 \ln 2}{8 \ln(1 + \sqrt{2}) - 3\sqrt{2} \ln 2 - 2\sqrt{2}} \\ & \times \frac{1}{G^2 \ln \Lambda} \frac{1}{m - m_0} \left(\frac{3M_0}{4\pi r_c^3} \right)^{-1} \left(\frac{GM_0}{2r_c} \right)^{3/2}. \end{aligned} \quad (22)$$

The numerical factor is about 0.3878.

The other extreme case is that the state $\beta = 1$ evolves directly into the state $\beta = \beta_{\text{fin}}$. These two states serve as the bases for the minor-component PDF, i.e., $f = (M_{\beta=1}/M)f_{\beta=1} + (M_{\beta=\beta_{\text{fin}}}/M)f_{\beta=\beta_{\text{fin}}} = (1 - \gamma)f_{\beta=1} + \gamma f_{\beta=\beta_{\text{fin}}}$ with $\gamma = 0$ at $t = 0$ ($0 \leq \gamma \leq 1$). The core mass M_c and the total energy E of the minor component are

$$M_c = (1 - \gamma) \frac{1}{2\sqrt{2}} M + \gamma M, \quad (23)$$

and

$$E = -(1 - \gamma) \frac{3\pi}{32} \frac{GM_0 M}{r_c} - \gamma \frac{GM_0 M}{r_c}. \quad (24)$$

Here we have made approximations for the state $\beta = \beta_{\text{fin}}$ taking account of the fact that the minor component is well concentrated in the core and has a negligible velocity dispersion. Then dM_c/dt is

obtained as $(dM_c/d\gamma)(d\gamma/dt) = (dM_c/d\gamma)(dE/dt)/(dE/d\gamma)$. Using their initial values, we obtain

$$\begin{aligned} \tau_{\text{ms}} = & \frac{3^2 \times 5 \times 7^2 \sqrt{\pi}}{2^{23}} \frac{19!!}{16!!} \frac{\Gamma(15/4)}{\Gamma(13/4)} \frac{32 - 3\pi}{2\sqrt{2} - 1} \\ & \times \frac{1}{G^2 \ln \Lambda} \frac{1}{m - m_0} \left(\frac{3M_0}{4\pi r_c^3} \right)^{-1} \left(\frac{GM_0}{2r_c} \right)^{3/2}. \end{aligned} \quad (25)$$

The numerical factor is about 0.6331. The present case has a somewhat larger timescale than the previous case because the effective energy difference is somewhat larger, i.e., $dE/d\gamma < dE/d\beta < 0$ while $dM_c/d\gamma \simeq dM_c/d\beta$ in the initial state.

Mass segregation in a real star cluster is expected to be intermediate between the above two cases. According to Fokker-Planck and Monte Carlo simulations of early stages of mass segregation (Inagaki & Wiyanto 1984; Fregeau et al. 2002), the ratio of velocity dispersions σ/σ_0 averaged over the whole cluster does not change appreciably, while that averaged over the core alone approaches steadily to the final value. The reason is that gravitational encounters occur preferentially among stars on orbits that pass through the dense core. Although this fact has been ignored in our present calculation, the behavior of σ/σ_0 would be reproduced if the minor-component PDF were a superposition of various states between those of $\beta = 1$ and $\beta = \beta_{\text{fin}}$ with the average β value increasing with the time.

Equations (22) and (25) have the same dependence on physical quantities. The values of the numerical factors are also similar to each other and well approximated to be 0.5 ± 0.1 or 2^{-1} . Thus the timescale of mass segregation is

$$\tau_{\text{ms}} = \frac{1}{2} \frac{1}{G^2 \ln \Lambda} \frac{1}{m - m_0} \left(\frac{3M_0}{4\pi r_c^3} \right)^{-1} \left(\frac{GM_0}{2r_c} \right)^{3/2}, \quad (26)$$

where $3M_0/4\pi r_c^3$ and $(GM_0/2r_c)^{1/2}$ correspond to the central mass density $\rho_0(0)$ and central velocity dispersion $\sigma_0(0)$ of the major component, respectively (eqs. [3] and [4]).

5. DISCUSSION

5.1. Comparison with Relevant Timescales

The timescale of mass segregation is greater than the core crossing timescale. Since the core radius is r_c and the central velocity dispersion is $(GM_0/2r_c)^{1/2}$, the core crossing timescale τ_{cc} is

$$\tau_{\text{cc}} = \left(\frac{2r_c^3}{GM_0} \right)^{1/2}. \quad (27)$$

Then equation (26) for the timescale of mass segregation τ_{ms} is rewritten as

$$\tau_{\text{ms}} = \frac{\pi}{6} \frac{m_0 N_0 / (m - m_0)}{\ln [m_0 N_0 / (m + m_0)]} \tau_{\text{cc}} \gg \tau_{\text{cc}}, \quad (28)$$

where we have assumed $m_0/m \simeq 10^{-1}$ – 10^{-2} and $N_0 \simeq 10^5$ – 10^6 as a practical situation. On the other hand, the timescale of mass segregation is less than the relaxation timescale τ_{rlx} of the major component, which is about $N_0/8 \ln N_0$ times the crossing timescale τ_{cc} (Binney & Tremaine 1987, p. 190):

$$\tau_{\text{rlx}} \simeq \frac{1}{8} \frac{N_0}{\ln N_0} \tau_{\text{cc}} > \tau_{\text{ms}}. \quad (29)$$

These results justifies our assumption that the star cluster is almost collisionless and the major-component PDF remains the same during mass segregation.

The timescale of mass segregation is close to the timescale of energy loss from the minor component, $\tau_{\text{el}} = E/(dE/dt)$ at $t = 0$. Using equations (10) and (16), we obtain

$$\tau_{\text{el}} = \frac{3^3 \times 5 \times 7^2 \pi \sqrt{\pi}}{2^{23}} \frac{19!!}{16!!} \frac{\Gamma(15/4)}{\Gamma(13/4)} \frac{1}{G^2 \ln \Lambda} \frac{1}{m - m_0} \left(\frac{3M_0}{4\pi r_c^3} \right)^{-1} \left(\frac{GM_0}{2r_c} \right)^{3/2}. \quad (30)$$

The dependence on physical quantities is the same as that of equation (26). This is because the energy loss from the minor component and the increase of its central concentration are of the same phenomenon. The numerical factor of equation (30) is about 0.4832, which happens to be in agreement with the numerical factor 2^{-1} of equation (26). Since this energy loss timescale is obtained without any assumption on the evolution of the minor-component PDF, it is a robust measure.

5.2. Comparison with Spitzer (1969)

Spitzer (1969) considered an isothermal cluster of uniform density and estimated the timescale of mass segregation as the timescale of approach toward energy equipartition between the major and minor components,² which was about m_0/m times the relaxation timescale τ_{rlx} of the major component. Using equation (29), we obtain the corresponding timescale in our model cluster:

$$\tau_{\text{ee (Spitzer)}} \simeq \frac{1}{8} \frac{m_0 N_0 / m}{\ln N_0} \tau_{\text{cc}}. \quad (31)$$

If $m \gg m_0$, this energy equipartition timescale is by definition equivalent to our energy loss timescale τ_{el} in equation (30). Since the numerical factor is 2^{-1} , the equation is rewritten as

$$\tau_{\text{el}} = \frac{\pi}{6} \frac{m_0 N_0 / m}{\ln(m_0 N_0 / m)} \tau_{\text{cc}}, \quad (32)$$

where we have replaced $m \pm m_0$ with m . Except for an unimportant difference in the argument of the logarithm, the timescales (31) and (32) have the same dependence on physical quantities. On the other hand, the numerical factors are different. This difference is not serious because there is

²Defined as the time it takes for the difference between $m\sigma^2$ and $m_0\sigma_0^2$ to decrease by a factor of e^{-1} .

ambiguity in the definition of the relaxation timescale. Our relaxation timescale (29) is based on the core crossing timescale τ_{cc} and might not be sufficiently large to yield the energy equipartition timescale for the whole cluster. Overall, although Spitzer (1969) considered a very idealized cluster, his timescale is expected to be valid for energy equipartition in a real star cluster at least as an order-of-magnitude estimation. This is also the case for mass segregation, which occurs during approach toward energy equipartition.

Spitzer (1969) also predicted that, if the total mass of the minor component is much less than that of the major component, $M \ll M_0$, the two components eventually achieve energy equipartition, $m\sigma^2 = m_0\sigma_0^2$. His prediction has been confirmed at least for the cluster core by Fokker-Planck and Monte Carlo simulations (Inagaki & Wiyanto 1984; Watters, Joshi, & Rasio 2000; Fregeau et al. 2002). Since it is required that both the major and minor components evolve to have thermal velocity distributions (§3), the energy equipartition is achieved at $t \gtrsim \tau_{\text{rlx}}$.³ This is beyond the limit of our model ($t < \tau_{\text{rlx}}$), where the cluster is assumed to be almost collisionless. The major component is kept to follow Plummer’s nonthermal model and hence does not achieve energy equipartition with the minor component, regardless of the total mass ratio M/M_0 . This fact is not serious to our estimation of the timescale of mass segregation, which is less than the relaxation timescale τ_{rlx} and has been defined at the initial time of the evolution $t = 0$ in equation (19).

5.3. Application to Observations

If the central mass density $\rho_0(0)$ and central velocity dispersion $\sigma_0(0)$ have values typical of globular clusters (Binney & Tremaine 1987, p. 26, see also p. 423 for the cutoff factor Λ), equation (26) is written as

$$\tau_{\text{ms}} = 3 \times 10^7 \text{ yr} \left(\frac{\ln \Lambda}{10} \right)^{-1} \left(\frac{m}{10 M_\odot} \right)^{-1} \left(\frac{\rho_0(0)}{10^4 M_\odot \text{ pc}^{-3}} \right)^{-1} \left(\frac{\sigma_0(0)}{10 \text{ km s}^{-1}} \right)^3. \quad (33)$$

Here we have assumed $m \gg m_0$ and replaced $m - m_0$ with m . If we instead use the typical values of the core radius r_c and central velocity dispersion $\sigma_0(0)$, both of which are observable quantities, equation (26) is rewritten as

$$\tau_{\text{ms}} = 2 \times 10^7 \text{ yr} \left(\frac{\ln \Lambda}{10} \right)^{-1} \left(\frac{m}{10 M_\odot} \right)^{-1} \left(\frac{r_c}{1 \text{ pc}} \right)^2 \left(\frac{\sigma_0(0)}{10 \text{ km s}^{-1}} \right). \quad (34)$$

Thus mass segregation in a typical star cluster has the timescale τ_{ms} of order 10^7 yr.

³This energy equipartition in the cluster core is not exact. The difference between $m\sigma^2$ and $m_0\sigma_0^2$ is very small but still existent ($\gtrsim 1\%$; Inagaki & Wiyanto 1984; Watters et al. 2000). The reason is that the velocity distributions do not become exactly thermal even in the cluster core.

The above timescale of mass segregation is used to discuss the origin of a very massive black hole found in the starburst galaxy M82 (Kaaret et al. 2001; Matsumoto et al. 2001; see also Matsushita et al. 2000). At the $2\ \mu\text{m}$ secondary peak, i.e., an active site of star formation, there is a source of compact X-ray emission. The observed strong variability implies that the source is an accreting black hole. The observed luminosity of $10^{41}\text{ ergs s}^{-1}$ implies that the mass is greater than $10^3 M_\odot$ if the emission is isotropic and its luminosity is below the Eddington limit. This black hole should have been formed via successive mergers of massive stars and black holes that had segregated into the core of a star cluster (Taniguchi et al. 2000). Since the $2\ \mu\text{m}$ secondary peak has a starburst age of about $1 \times 10^7\text{ yr}$ (Satyapal et al. 1997), the mass segregation should have occurred well within this short duration. The timescales given in equations (33) and (34) are too long. We suspect that the star cluster is exceptionally dense and compact. For example, the young star cluster R136 in the Large Magellanic Cloud has $\rho_0(0) \simeq 10^6 M_\odot \text{ pc}^{-3}$, $\sigma_0(0) \simeq 10^1 \text{ km s}^{-1}$, and $r_c \simeq 10^{-2}\text{--}10^{-1}\text{ pc}$ (Portegies Zwart et al. 1999), which yields $\tau_{\text{ms}} \simeq 10^5\text{ yr}$. In fact, only one black hole heavier than $10^3 M_\odot$ has been found among more than 100 young star clusters of M82. The objects that had segregated into the cluster core should have merged in a runaway manner (Quinlan & Shapiro 1989; Portegies Zwart et al. 1999; Mouri & Taniguchi 2002) because the merger probability is higher for more massive objects.

This work has been supported in part by the Japanese Ministry of Education, Science, and Culture under grants 10044052 and 10304013. We are grateful to N. Kanaeda and the referee for helpful comments.

REFERENCES

- Binney, J., & Tremaine, S. 1987, *Galactic Dynamics* (Princeton: Princeton Univ. Press)
- Chapman, S., & Cowling, T. G. 1970, *The Mathematical Theory of Non-Uniform Gases* (3rd ed.; Cambridge: Cambridge Univ. Press)
- Chatterjee, P., Hernquist, L., & Loeb, A. 2002, *Phys. Rev. Lett.*, 88, 121103
- Fregeau, J. M., Joshi, K. J., Portegies Zwart, S. F., & Rasio, F. A. 2002, *ApJ*, 570, 171
- Inagaki, S., & Wiyanto, P. 1984, *PASJ*, 36, 391
- Kaaret, P., et al. 2001, *MNRAS*, 321, L29
- Matsumoto, H., et al. 2001, *ApJ*, 547, L25
- Matsushita, S., et al. 2000, *ApJ*, 545, L107
- Mouri, H., & Taniguchi, Y. 2002, *ApJ*, 566, L17
- Portegies Zwart, S. F., Makino, J., McMillan, S. L. W., & Hut, P. 1999, *A&A*, 348, 117
- Quinlan, G. D. & Shapiro, S. L. 1989, *ApJ*, 343, 725

Satyapal, S., et al. 1997, ApJ, 483, 148

Spitzer, L., Jr. 1969, ApJ, 158, L139

Taniguchi, Y., Shioya, Y., Tsuru, T. G., & Ikeuchi, S. 2000, PASJ, 52, 533

Tremaine, S. D., Ostriker, J. P., & Spitzer, L., Jr. 1975, ApJ, 196, 407

Watters, W. A., Joshi, K. J., & Rasio, F. A. 2000, ApJ, 539, 331

FIGURE CAPTIONS

Fig. 1.— *Solid lines:* Numerical values of F_1 , F_2 , F_3 , F_4 , F_5 , and F_6 . *Dotted lines:* Numerical values of F_1F_2 , F_1F_3 , $F_1F_2F_4$, $F_1F_2F_4F_5$, and F_1F_6 . The abscissa is the parameter β .

Fig. 2.— *Solid lines:* Radial profiles of the mass density ρ_β in equation (8) for $\beta = 1$, 10, and 100. *Dotted lines:* Radial profiles of the squared velocity dispersion σ_β^2 in equation (9). The abscissa is the radius r normalized by the core radius r_c . For convenience, we have assumed $3M/4\pi r_c^3 = GM_0/2r_c = 1$.

Table 1. NUMERICAL VALUES OF THE FACTORS F_1 – F_6

β	F_1	F_2	F_3	F_4	F_5	F_6
1	1.000	1.000	1.000	1.000	1.000	1.000
2	3.875	2.294	0.3469	0.6177	0.5424	0.5480
3	7.761	3.902	0.1881	0.4292	0.3519	0.3325
5	17.81	7.868	0.08728	0.2508	0.1914	0.1573
10	52.71	21.15	0.03083	0.1081	0.07705	0.05366
20	152.4	58.30	0.01090	0.04263	0.02921	0.01856
30	282.1	106.2	0.005933	0.02412	0.01628	0.01003
50	610.4	226.8	0.002757	0.01157	0.007717	0.004633
100	1734.	638.1	0.0009748	0.004189	0.002769	0.001631